Comparison of NCEP performance specifications for triglycerides, HDL-, and LDL-cholesterol with operating specifications based on NCEP clinical and analytical goals

PATRICIA C. FALLEST-STROBL,1 ELIN OLAIFSDOTTIR,2 DONALD A. WIEBE,1 and JAMES O. WESTGARD1∗

The National Cholesterol Education Program (NCEP) performance specifications for methods that measure triglycerides, HDL-cholesterol, and LDL-cholesterol have been evaluated by deriving operating specifications from the NCEP analytical total error requirements and the clinical requirements for interpretation of the tests. We determined the maximum imprecision and inaccuracy that would be allowable to control routine methods with commonly used single and multirule quality-control procedures having 2 and 4 control measurements per run, and then compared these estimates with the NCEP guidelines. The NCEP imprecision specifications meet the operating imprecision necessary to assure meeting the NCEP clinical quality requirements for triglycerides and HDL-cholesterol but not for LDL-cholesterol. More importantly, the NCEP imprecision specifications are not adequate to assure meeting the NCEP analytical total error requirements for any of these three tests. Our findings indicate that the NCEP recommendations fail to adequately consider the quality-control requirements necessary to detect medically important systematic errors.

Three different types of performance criteria have been established for lipid tests performed in healthcare laboratories. For example, cholesterol tests are regulated by the Clinical Laboratory Improvement Act (CLIA-88), which requires laboratories to provide test values that are correct within 10% of target values for proficiency testing materials (a total error criterion). The US National Cholesterol Education Program (NCEP)3 provides physicians with explicit guidelines to interpret a patient’s cholesterol values (a clinical criterion) and also provides laboratories with specifications for the imprecision and inaccuracy of their routine cholesterol methods (analytical criteria). Guidelines published by NCEP recommended that by 1992 the performance of cholesterol tests should have a 3.0% or less CV (or imprecision) and 3.0% or less bias (inaccuracy) [1]. We evaluated these recommendations in earlier studies [2–4] and expressed concern that QC issues had not been adequately considered, in that a CV of 2.4% or less was actually required to assure that routine methods would achieve the performance required for the clinical use of the test.

Because cholesterol is a key risk factor for cardiovascular disease, the importance for establishing both clinical and analytical goals cannot be argued. However, these goals are interdependent (i.e., a method’s analytical performance will often determine the clinical usefulness of a given test and, likewise, clinical demands often dictate the stringency required of a method’s analytical performance). The two groups that established these guidelines, the Adult Treatment Panel and the Laboratory Standardization Panel, failed to interconnect the different performance criteria and address them together. Nonetheless, the primary goals of NCEP have been overwhelmingly

1 Department of Pathology and Laboratory Medicine, University of Wisconsin Medical School, Madison, WI 53792.
2 Department of Clinical Biochemistry, University Hospital, Reykjavik, Iceland.
*Author for correspondence. Fax (608)263-1568; e-mail jo.westgard@hosp.wisc.edu.
Received February 25, 1997; revision accepted June 26, 1997.

3 Nonstandard abbreviations: NCEP, National Cholesterol Education Program; OPSpecs, operating specifications for allowable imprecision (CV), allowable inaccuracy (bias), and necessary QC; TEa, total allowable analytical error; Dint, clinical decision interval; swsub, within-subject biological variation; TRIG, triglycerides; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol.
achieved. Family physicians appropriately treat hypercholesterolemia in patients with aggressive therapy to lower cholesterol to lessen their cardiovascular risk. Laboratories have improved their performance and are providing more-accurate cholesterol test values. The laboratory success was driven in part by manufacturers that provided improved instruments and techniques designed to meet the NCEP guidelines.

Cholesterol, however, is only one of several lipid components of clinical importance for which appropriate guidelines are needed. In 1995, the NCEP Laboratory Standardization Panel published recommendations of performance specifications for triglycerides (TRIG), HDL-cholesterol (HDL-C), and LDL-cholesterol (LDL-C) [5–7]. These recommendations define specifications for the imprecision and inaccuracy of the respective methods, the total analytical error that is allowable, and the medical decision values. However, lack of guidance for the routine operation and QC of these methods prompted us to evaluate the operating specifications necessary to assure that the NCEP clinical and analytical goals could be achieved in routine lipid testing, taking into account the QC performance that would be necessary to detect analytical runs in which method performance is unstable. Detailed specifications for QC are essential for these methods, which are prone to many analytical variables that could affect routine laboratory performance and result in inappropriate patient data transmitted to the patient.

Materials and Methods
OPSpecs charts [8] were prepared by using the QC Validator computer program [9] (Westgard Quality Corp., Ogunquit, ME). The principles of the OPSpecs methodology have been described in detail elsewhere [10]. OPSpecs charts plot the allowable inaccuracy on the y-axis vs the allowable imprecision on the x-axis for a defined quality requirement and a specified requirement of error detection, or analytical quality assurance. The observed imprecision and inaccuracy of an individual method can be plotted to locate the operating point of the method. Appropriate QC procedures are provided by the control rules and number of control measurements per run (N), the operating limits of which are above the method’s observed operating point [11]. Estimates of maximum allowable imprecision can be obtained from the x-intercepts of the operating limits [12].

In this study, OPSpecs charts were prepared to show the imprecision and inaccuracy that are allowable for common QC procedures with N = 2 to 4, 90% analytical quality assurance, and the stated NCEP analytical total error requirements and the clinical decision interval (D_int) requirements implied by the guidelines for test interpretation [5–7]. We calculated D_int for each of the tests according to the medical decision values for treatment of patients as recommended by the Adult Treatment Panel. For TRIG, concentrations for healthy subjects are <2000 mg/L and high concentrations are ≥4000 mg/L; therefore, the D_int is 100%: (2000/2000) × 100%. For HDL-C, desirable concentrations are >600 mg/L and major risk concentrations are <350 mg/L; therefore, D_int is 42%: (250/600) × 100%. For LDL-C, desirable values are <1300 mg/L and high-risk concentrations are ≥1600 mg/L; therefore, D_int is 23.1%: (300/1300) × 100%. Estimates of the within-subject biological variation were taken as 23.7% for TRIG [6], 7.5% for HDL-C [7], and 8.2% for LDL-C [5], as recommended in the NCEP documents. The range of allowable imprecision (operating imprecision in Table 1) was determined at the bias stated for each test.

Results
Clinical quality requirements. The OPSpecs charts for the NCEP clinical quality requirements for TRIG, HDL-C, and LDL-C are shown in Fig. 1. For TRIG, the operating point that represents the NCEP-recommended specifications (inaccuracy of 5%, imprecision of 5%) is well below the allowable range of imprecision (14.0–22.5%) at the stated value of 5% for inaccuracy, as denoted by the double-headed arrow in Fig. 1A. Therefore, the TRIG specifications for imprecision and inaccuracy will assure the quality defined by the NCEP clinical requirement (D_int = 100%; s_wsub = 23.7%). The operating point for HDL-C falls within the allowable range of imprecision (5.0–7.8%) for the control rules with N = 2–4 (Fig. 1B); thus, the HDL-C specifications for accuracy (10.0%) and precision (6.0%) will also assure the quality defined by the NCEP clinical requirement (D_int = 42.0%, s_wsub = 7.5%). However, the LDL-C operating point of 4.0% imprecision and 4.0% inaccuracy exceeds the allowable range of imprecision (1.5–2.6%), as shown in Fig. 1C. Therefore, the NCEP

<table>
<thead>
<tr>
<th>Test and requirements</th>
<th>NCEP bias</th>
<th>NCEP CV, %</th>
<th>Operating imprecision</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRIG</td>
<td>5.0</td>
<td>5.0</td>
<td>14.0–22.5</td>
</tr>
<tr>
<td>D_int = 100%, s_wsub = 23.7%</td>
<td>0</td>
<td>0</td>
<td>15.5–24.0</td>
</tr>
<tr>
<td>HDL-C</td>
<td>10.0</td>
<td>6.0</td>
<td>5.0–7.8</td>
</tr>
<tr>
<td>D_int = 42.0%, s_wsub = 7.5%</td>
<td>0</td>
<td>0</td>
<td>7.0–11.0</td>
</tr>
<tr>
<td>LDL-C</td>
<td>4.0</td>
<td>4.0*</td>
<td>1.5–2.6</td>
</tr>
<tr>
<td>D_int = 23.1%, s_wsub = 8.2%</td>
<td>0</td>
<td>0</td>
<td>2.5–4.2</td>
</tr>
</tbody>
</table>

Table 1. Comparison of NCEP clinical and analytical requirements with operating imprecision derived from OPSpecs charts.

Clinical decision interval, s_wsub = within-subject biological variation, TE_a = total allowable error, bias = allowable inaccuracy, CV = coefficient of variation or allowable imprecision. * = NCEP recommendation exceeds the operating imprecision.
method performance specifications for LDL-C will not assure the quality defined by the NCEP clinical requirement ($D_{\text{int}} = 23.1\%$, $s_{\text{sub}} = 8.2\%$).

**Analytical quality requirements.** The operating point for TRIG (inaccuracy 5%, imprecision 5%) exceeds the allowable range of imprecision (2.2–3.2%) for analytical quality requirements ($T_E = 15.0\%$), as depicted by the OPSpecs chart in Fig. 2A. Therefore, common QC procedures with $N = 2–4$ will not be able to detect unstable performance. Similarly, the OPSpecs charts for HDL-C and LDL-C (Fig. 2B and C) show operating points that also exceed the allowable range of imprecision for detection of unstable performance. Therefore, the analytical quality requirements for precision of test methods for TRIG, HDL-C, and LDL-C will not assure the analytical quality defined by NCEP.

NCEP’s recommendations for HDL-C [7] also included performance specifications ($T_E = 13\%$, $CV = 4.0\%$, bias = 5.0%) to be achieved by 1998. Although these specifications are more stringent than the current criteria for HDL-C, analytical quality requirements are still not satisfied. OPSpecs charts prepared for use of common QC procedures (not shown) indicate that the allowable imprecision for HDL-C methods still needs to be $<4.0\%$ (within the range 1.5–2.3%) at a bias of 5.0% and $T_E = 13.0\%$.

Table 1 summarizes the ranges of allowable imprecision derived from OPSpecs charts, based on the NCEP clinical and analytical quality requirements for all of these tests. Also shown are the allowable ranges of imprecision for each test if bias were eliminated as a consideration (bias = 0). In this scenario, the allowable ranges of imprecision are increased for all of the tests, but NCEP analytical quality requirements are still not assured for TRIG or LDL-C (Table 1).

**Discussion**

Methods for determination of TRIG, HDL-C, and LDL-C are complex and can be problematic, which means it is important to select optimal QC procedures capable of detecting unstable performance. TRIG measurements are often automated, whereas HDL-C and LDL-C methods usually include manual steps in test protocols. LDL-C test results may be the most problematic because common practice is to report estimated LDL-C by calculation with the Friedewald equation [estimated LDL = total cholesterol $- HDL - (TRIG/5)$]. Therefore, each LDL-C result relies on accurate and precise measurements of three other lipids: cholesterol, HDL-C, and TRIG. The ability to detect unstable performance in routine laboratory testing for lipids through use of statistical QC procedures is crucial to avoid further propagation of error in generating LDL-C test results.

Each laboratory should pay special attention to the operating specifications needed to assure that the desired analytical or clinical quality is achieved in routine service. OPSpecs charts provide a graphical display of the impre-
cision and inaccuracy that are allowable when different QC procedures are used. In this approach, statistical QC is factored in as a component of the error budget for a test, thus providing a more realistic assessment of the imprecision and inaccuracy limits that are necessary if the QC procedure is to detect medically important errors. OP-Specs charts, therefore, are a very useful tool for defining the specifications needed for lipid tests in a routine service laboratory.

The NCEP analytical recommendations for imprecision and inaccuracy seem to be based on the method acceptance criterion of $TE_a = \text{bias} + 2 \times s$. For methods that just meet this criterion, commonly used control procedures are not expected to provide reliable detection of medically important errors [13]. Even if bias could be eliminated from lipid test methods, the derived data for operating precision (when bias = 0) show that NCEP-recommended precision goals for TRIG and LDL-C are not satisfactory to meet the stated analytical total error requirements. Laboratories should select methods for these tests for which operating imprecision is within the ranges indicated in Table 1 to assure adequate detection of unstable performance. In general, further reducing by a factor of 2 the limits set by NCEP for allowable imprecision would be desirable to provide more-controllable testing processes.

In conclusion, improvements in lipid test performance and interpretation have continued over the years because of efforts by NCEP, but our findings indicate that additional guidelines are needed for statistical QC to assure detection of unstable performance. Although NCEP’s 1998 goals for HDL-C are a step in the right direction, they will still fail to provide a high degree of analytical quality assurance. In the future, development of performance specifications for analytical methods should take into account the performance characteristics of the statistical QC procedures that are expected to be used in monitoring routine laboratory performance. The NCEP Adult Treatment Panel and Laboratory Standardization Panel independently generated clinical and analytical goals that have conflicting or unobtainable targets. Clearly, laboratorians will need guidance to establish acceptable performance standards for these routine, yet demanding lipid tests. However, if performance goals result in unrealistic objectives, the final outcome could be misuse or inappropriate handling of proficiency testing material to ensure that the analytical goals are met, while patients’ needs are ignored. These activities must be coordinated with clearer QC operations for laboratories to properly manage these tests.

The Department of Pathology and Laboratory Medicine, University of Wisconsin Medical School, provided partial support for E.O. during a sabbatical period and full support for P.F.-S. during a Clinical Chemistry Fellowship.
References